

cm Newman
100548

CGNE 119-2 US

PLANT VDE GENES AND METHODS RELATED THERETO

Field of the Invention

This invention relates to genes encoding plant violaxanthin de-epoxidase (vde) and methods of use related to the protein and the nucleic acid sequences. The invention is exemplified by methods of causing increased expression or decreased expression of plant vde genes in plants. Included are plants produced by the method.

INTRODUCTION

Background

Plant carotenoids are found in the membranes of chloroplasts and chromoplasts. They are instrumental in the photoprotective mechanisms of plants. Also, plant carotenoids have significant dietary implications. Thus, from an agronomic as well as a nutritional standpoint, study of the plant carotenoids and the enzymes involved in the biosynthesis of carotenoids is of interest.

Of particular interest are the late stages of the carotenoid biosynthetic pathway in plants, the xanthophyll cycle and its importance in photoregulation of photosynthesis. Photosynthesis is the process that enable plants to use light energy for growth and development. Thus, the availability of light of appropriate quality and quantity (photosynthetically active radiation or "PAR") is critical for plant growth and development. Ironically, light can also damage plants because plants have limited capacity to use light. When light intensity exceeds this capacity, irreversible damage can occur.

Plants have developed various mechanisms to cope with excess light such as varying leaf orientation or developing reflective surfaces. Such mechanisms appear to be specialized phenotypic strategies that are limited to certain types of

plants. One mechanism that is apparently used by all plants examined so far is the dissipation of excess energy as heat in the antenna (light absorbing structures) of the photosynthetic apparatus. Most of the excess energy is discarded as heat by a complex feed-back regulatory system that involves the transthylakoid ΔpH and formation of antheraxanthin and zeaxanthin catalyzed by violaxanthin de-epoxidase (vde) in the xanthophyll cycle. This system, termed energy dependent non-radiative energy dissipation or non-photochemical fluorescence quenching, reduces the quantum efficiency of photosystem II (PSII), helping to prevent PSII over reduction and photoinhibitory damage. In effect, this system provides a means to dump excess energy before it can damage the photosynthetic apparatus. The system has a wide dynamic range, both qualitatively and quantitatively, which enables it to function effectively over a wide-range of environmental conditions.

The ability to manipulate aspects of the xanthophyll cycle through genetic engineering techniques would permit the rapid introduction of improved plant varieties. However, it has been difficult to obtain purified fractions of the enzymes involved in the pathway and, prior to this invention, the corresponding genes have not been cloned.

SUMMARY OF THE INVENTION

DNA sequences encoding plant vde enzymes are provided herein. The sequences may be joined to heterologous DNA sequences for use as probes and in DNA constructs to modify the genotype of a host organism. DNA constructs and methods are provided to modify a host cell phenotype by altering the amount of photoprotection enzyme present in the host cell. In plastid containing host cells, zeaxanthin levels and sensitivity to light can be modified through alterations in the level of vde enzymes.

For example, over expression of vde is expected to increase the tolerance of plants to high light, drought and temperature stress (stress conditions exacerbate the condition

of excess light). Also, plants that are not currently tolerant to high light or low temperatures are expected to become more tolerant to these stresses. Plants that are better adapted to light stress are expected to be more productive and/or more resistant to disease. Alternatively, the under expression, or inhibition of vde activity is expected to increase photosynthetic efficiency under low light. The growing range of plants, crops, trees and ornamentals, could thus be modified.

Specific plant vde's are described. In particular, a 55 kD lettuce vde having the cDNA sequence and deduced amino acid sequence as shown in Fig. 1, a tobacco vde having the cDNA sequence and deduced amino acid sequence as shown in Fig. 2, and an *Arabidopsis* vde having the cDNA sequence and deduced amino acid sequence as shown in Fig. 3, are described. Figure 4 provides a comparison at the amino acid level of the proteins of Figures 1-3. In this amino acid sequence comparison the transit peptides for the three sequences are boxed. Identical amino acids are denoted by a hyphen. Gaps inserted to optimize sequence alignments are denoted with a period. The thirteen highly conserved cysteine residues are denoted with an asterisk.

Figure 5 is a comparison of the identity and similarity of pre-protein and mature protein vde. As can be seen from Figure 5, diverse vde's have sequences with about 75% sequence identity with one another at the amino acid level. Thus, vde sequences having at least about 75% homology to amino acid sequences in Fig.1, Fig.2 or Fig. 3 are also contemplated hereunder.

Nucleic acid sequences encoding a plant vde having at least about 60% sequence identity, and more preferably at least about 70% sequence identity, with the sequences in Figs. 1, 2 or 3, and are likewise contemplated herein. For instance, a comparison of tobacco and lettuce vde nucleic acid sequences give 76% identity, excluding the transit peptides. A high degree of sequence identity at the N-terminus is particularly preferred. Other related plant photoregulatory

sequences having high degrees of similarity with fragments of the vde sequences shown are also contemplated.

In a different aspect of this invention, nucleic acid sequences related to the exemplified lettuce, tobacco and arabidopsis vde sequences of this invention are described with details regarding methods to obtain such sequences from a variety of sources and their use. In addition, cDNA sequences encoding mature vde's are given as well as transit peptides, mRNA, genomic plant vdes, and plant vde regulatory regions.

In a further aspect of this invention, methods of producing vde in host cells are described. In plastid containing cells, modifications in the xanthophyll cycle, particularly in the ratio of violaxanthin as to zeaxanthin are contemplated via increased production of vde or decreased production of vde. This will have applications in the increased feed value of plants. Zeaxanthin levels are important to crops such as alfalfa whose value in part is due to xanthophyll content.

Results from studies of transgenic plants demonstrates that xanthophyll-mediated energy dissipation in LHCII apparently protects PSII against the potentially damaging effects of high light. This protection is induced by the combined effects of a thylakoid ΔpH and the presence of zeaxanthin and antheraxanthin formed by violaxanthin de-epoxidase (vde) activity.

DESCRIPTION OF THE FIGURES

FIG. 1 cDNA sequence for romaine lettuce vde and deduced polypeptide sequence. (SEQ ID NO: 1 and 2) The underlined sequences are those determined from peptide sequencing of purified lettuce vde. The polypeptide sequence begins at the first methionine of the open reading frame and is preceded by three termination codons in the same reading frame.

FIG. 2 cDNA sequence for tobacco (*Nicotiana tabacum* cv. Xanthi) vde and deduced polypeptide sequence. (SEQ ID NO: 3 and 4)

FIG. 3 cDNA sequence for *Arabidopsis thaliana* (var. columbia) vde and deduced polypeptide sequence. (SEQ ID NO: 5 and 6)

a FIG. 4 provides a comparison of the amino acid sequences of the proteins of Figures 1-3.

FIG. 5 shows the percent similarity between the the proteins of Figures 1-3.

FIG. 6 provides a comparison of hydropathy profiles for the vdes of three species.

FIG. 7 provides a time-course comparison of effects of expressed vde.

ms. D' ~~FIG. 8 is a table showing the results of pigment analysis of leaves of control and 18 vde-antisense tobacco plants (TAS-#).~~

FIG. 9 shows the results of a control plant extraction for vde.

FIG. 10 shows the results of extraction for vde in an antisense vde plant.

DETAILED DESCRIPTION OF THE INVENTION

A plant violaxanthin de-epoxidase or "vde" of this invention includes any sequence of amino acids, such as a protein, polypeptide or peptide, obtainable from a plant source, which demonstrates the ability to catalyze the production of zeaxanthin from violaxanthin under plant enzyme reactive conditions. By "enzyme reactive conditions" is meant that any necessary conditions that are available in an environment (i.e., such factors as temperature, pH, lack of inhibiting substances) which will permit the enzyme to function.

By "plant" is meant any plastid-containing organism. A "higher plant" shall mean any differentiated, multi-cellular plastid-containing organism. Of particular interest are plant vde's from angiosperms, both dicotyledonous and monocotyledonous plants.

In this invention, the cDNA sequence of a lettuce (Fig. 1), tobacco (Fig. 2) and *Arabidopsis* (Fig. 3) vde gene are provided. Transit peptide regions are identified in Fig. 4. From these sequences, genomic sequences may be obtained and the corresponding transcriptional and translational regulatory

regions determined. Also, using the lettuce and/or tobacco sequences provided, vde genes from other sources may be obtained. In particular, it is found that the N-terminal regions of the lettuce, tobacco, *Arabidopsis* and spinach proteins are conserved and therefore, an N-terminal peptide such as "VDALKTCACLLK" ^(see also no. 7) will find particular use in obtaining related sequences.

Constructs for use in the methods may include several forms, depending upon the intended use of the construct. Thus, the constructs include vectors, transcriptional cassettes, expression cassettes and plasmids. The transcriptional and translational initiation region (also sometimes referred to as a "promoter") preferably comprises a transcriptional initiation regulatory region and a translational initiation regulatory region of untranslated 5' sequences, "ribosome" binding sites, responsible for binding mRNA to ribosomes and translational initiation. It is preferred that all of the transcriptional and translational functional elements of the initiation control region are derived from or obtainable from the same gene. In some embodiments, the promoter will be modified by the addition of sequences, such as enhancers, or deletions of nonessential and/or undesired sequences. By "obtainable" is intended a promoter having a DNA sequence sufficiently similar to that of a native promoter to provide for the desired specificity of transcription of a DNA sequence of interest. It includes natural and synthetic sequences as well as sequences which may be a combination of synthetic and natural sequences.

A transcriptional cassette for transcription of a nucleotide sequence of interest will include in the direction of transcription, a transcription initiation region and optionally a translational initiation region, a DNA sequence of interest, and a transcriptional and optionally translational termination region functional in the host cell of interest. When the cassette provides for the transcription and translation of a DNA sequence it is considered an expression cassette. One or more introns may also be present.

Other sequences may also be present, including those encoding transit peptides.

The use of amino acid sequences from vde peptides to obtain nucleic acid sequences which encode lettuce vde is described herein. For example, synthetic oligonucleotides are prepared which correspond to the vde peptide sequences. The oligonucleotides are used as primers in polymerase chain reaction (PCR) techniques to obtain partial DNA sequence of vde genes. The partial sequences so obtained are then used as probes to obtain vde clones from a gene library prepared from lettuce tissue. Alternatively, where oligonucleotides of low degeneracy can be prepared from particular vde peptides, such probes may be used directly to screen gene libraries for vde gene sequences. In particular, screening of cDNA libraries in phage vectors is useful in such methods due to lower levels of background hybridization.

A nucleic acid sequence of a plant vde of this invention may be a DNA or RNA sequence, derived from genomic DNA, cDNA, mRNA, or may be synthesized in whole or in part. The gene sequences may be cloned, for example, by isolating genomic DNA from an appropriate source, and amplifying and cloning the sequence of interest using a polymerase chain reaction (PCR). Alternatively, the gene sequences may be synthesized, either completely or in part, especially where it is desirable to provide plant-preferred sequences. Thus, all or a portion of the desired structural gene (that portion of the gene which encodes the vde protein) may be synthesized using codons preferred by a selected host. Host-preferred codons may be determined, for example, from the codons used most frequently in the proteins expressed in a desired host species.

One skilled in the art will readily recognize that antibody preparations, nucleic acid probes (DNA and RNA) and the like may be prepared and used to screen and recover "homologous" or "related" vde's from a variety of plant sources. Homologous sequences are found when there is an identity of sequence, which may be determined upon comparison of sequence information, nucleic acid or amino acid, or

through hybridization reactions between a known vde and a candidate source. Conservative changes, such as Glu/Asp, Val/Ile, Ser/Thr, Arg/Lys and Gln/Asn may also be considered in determining sequence homology. Amino acid sequences are considered homologous by as little as 25% sequence identity between the two complete mature proteins. (See generally, Doolittle, R.F., *OF URFS and ORFS* (University Science Books, CA, 1986.)

Thus, other plant vde's may be obtained from the specific exemplified lettuce, tobacco and *Arabidopsis* sequences provided herein. Furthermore, it will be apparent that one can obtain natural and synthetic plant vde's, including modified amino acid sequences and starting materials for synthetic-protein modeling from the exemplified plant vde's and from plant vde's which are obtained through the use of such exemplified sequences. Modified amino acid sequences include sequences which have been mutated, truncated, increased and the like, whether such sequences were partially or wholly synthesized. Sequences which are actually purified from plant preparations or are identical or encode identical proteins thereto, regardless of the method used to obtain the protein or sequence, are equally considered naturally derived.

Typically, a plant vde sequence obtainable from the use of nucleic acid probes will show 60-70% sequence identity between the target vde sequence and the encoding sequence used as a probe. However, lengthy sequences with as little as 50-60% sequence identity may also be obtained. The nucleic acid probes may be a lengthy fragment of the nucleic acid sequence, or may also be a shorter, oligonucleotide probe. When longer nucleic acid fragments are employed as probes (greater than about 100 bp), one may screen at lower stringencies in order to obtain sequences from the target sample which have 20-50% deviation (i.e., 50-80% sequence homology) from the sequences used as probe. Oligonucleotide probes can be considerably shorter than the entire nucleic acid sequence encoding a vde enzyme, but should be at least about 10, preferably at least about 15, and more preferably at least about 20 nucleotides.

A higher degree of sequence identity is desired when shorter regions are used as opposed to longer regions. It may thus be desirable to identify regions of highly conserved amino acid sequence to design oligonucleotide probes for detecting and recovering other related vde genes. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be identified. (See, Gould, et al., *PNAS USA* (1989) 86:1934-1938.)

To determine if a related gene may be isolated by hybridization with a given sequence, the sequence is labeled to allow detection, typically using radioactivity, although other methods are available. The labeled probe is added to a hybridization solution, and incubated with filters containing the desired nucleic acids, either Northern or Southern blots (to screen desired sources for homology), or the filters containing cDNA or genomic clones to be screened. Hybridization and washing conditions may be varied to optimize the hybridization of the probe to the sequences of interest. Lower temperatures and higher salt concentrations allow for hybridization of more distantly related sequences (low stringency). If background hybridization is a problem under low stringency conditions, the temperature can be raised either in the hybridization or washing steps and/or salt content lowered to improve detection of the specific hybridizing sequence. Hybridization and washing temperatures can be adjusted based on the estimated melting temperature of the probe as discussed in Beltz, et al. (*Methods in Enzymology* (1983) 100:266-285).

A useful probe and appropriate hybridization and washing conditions having been identified as described above; cDNA or genomic libraries are screened using the labeled sequences and optimized conditions. The libraries are first plated onto a solid agar medium, and the DNA lifted to an appropriate membrane, usually nitrocellulose or nylon filters. These filters are then hybridized with the labeled probe and washed as discussed above to identify clones containing the related sequences. When a genomic library is used, one or more

sequences may be identified providing both the coding region and the transcriptional regulatory elements of the vde gene from such plant source.

For immunological screening, antibodies to the vde protein can be prepared by injecting rabbits or mice with the protein purified from the original plant source or expressed from a host cell, such methods of preparing antibodies being well known to those in the art. Either monoclonal or polyclonal antibodies can be produced, although typically polyclonal antibodies are more useful for gene isolation. Western analysis may be conducted to determine that a related protein is present in a crude extract of the desired plant species, as determined by cross-reaction with the antibodies to the vde. When cross-reactivity is observed, genes encoding the related proteins are isolated by screening expression libraries representing the desired plant species. Expression libraries can be constructed in a variety of commercially available vectors, including lambda gt11, as described in Maniatis, et al. (*supra*).

All plants studied to date utilize the xanthophyll cycle, and thus any given plant species can be considered as a source of additional vde proteins.

The nucleic acid sequences associated with plant vde proteins will find many uses. For example, recombinant constructs can be prepared which can be used as probes or will provide for expression of the vde protein in host cells to produce a ready source of the enzyme. Other useful applications may be found when the host cell is a plant host cell, either *in vitro* or *in vivo*. For example, by increasing the amount of a respective vde available to the plant xanthophyll cycle, an increased percentage of zeaxanthin may be obtained. In a like manner, for some applications it may be desired to decrease the amount of vde endogenously expressed in a plant cell by anti-sense or some other reducing technology such as co-suppression. For example, to improve photosynthetic efficiency of a plant under low light, decreased expression of a vde may be desired.

Thus, depending upon the intended use, the constructs may contain the sequence which encodes the entire vde protein, or a portion thereof. For example, where antisense inhibition of a given vde protein is desired, the entire vde sequence is not required. Furthermore, where vde constructs are intended for use as probes, it may be advantageous to prepare constructs containing only a particular portion of an vde encoding sequence, for example a sequence which is discovered to encode a highly conserved vde region.

As discussed above, nucleic acid sequence encoding a plant vde of this invention may include genomic, cDNA or mRNA sequence. By "encoding" is meant that the sequence corresponds to a particular amino acid sequence either in a sense or anti-sense orientation. By "extrachromosomal" is meant that the sequence is outside of the plant genome of which it is naturally associated. By "recombinant" is meant that the sequence contains a genetically engineered modification through manipulation via mutagenesis, restriction enzymes, and the like.

A cDNA sequence may or may not contain pre-processing sequences, such as transit peptide sequences or targeting sequences to facilitate delivery of the vde protein to a given organelle or membrane location. The use of any such precursor vde DNA sequences is preferred for uses in plant cell expression. A genomic vde sequence may contain the transcription and translation initiation regions, introns, and/or transcript termination regions of the plant vde, which sequences may be used in a variety of DNA constructs, with or without the vde structural gene. Thus, nucleic acid sequences corresponding to the plant vde of this invention may also provide signal sequences useful to direct protein delivery into a particular organelle or membrane location, 5' upstream non-coding regulatory regions (promoters) having useful tissue and timing profiles, 3' downstream non-coding regulatory region useful as transcriptional and translational regulatory regions and may lend insight into other features of the gene.

Once the desired plant vde nucleic acid sequence is obtained, it may be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions may be subjected to resection, mutagenesis, etc. Thus, transitions, transversions, deletions, and insertions may be performed on the naturally occurring sequence. In addition, all or part of the sequence may be synthesized. In the structural gene, one or more codons may be modified to provide for a modified amino acid sequence, or one or more codon mutations may be introduced to provide for a convenient restriction site or other purpose involved with construction or expression. The structural gene may be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like.

The nucleic acid or amino acid sequences encoding a plant vde of this invention may be combined with other non-native, or "heterologous", sequences in a variety of ways. By "heterologous" sequences is meant any sequence which is not naturally found joined to the plant vde, including, for example, combinations of nucleic acid sequences from the same plant which are not naturally found joined together.

The DNA sequence encoding a plant vde of this invention may be employed in conjunction with all or part of the gene sequences normally associated with the vde. In its component parts, a DNA sequence encoding vde is combined in a DNA construct having, in the 5' to 3' direction of transcription, a transcription initiation control region capable of promoting transcription and translation in a host cell, the DNA sequence encoding plant vde and a transcription and translation termination region.

Potential host cells include both prokaryotic and eukaryotic cells. A host cell may be unicellular or found in a multicellular differentiated or undifferentiated organism depending upon the intended use. Cells of this invention may be distinguished by having a plant vde foreign to the wild-type cell present therein, for example, by having a

recombinant nucleic acid construct encoding a plant vde therein.

Depending upon the host, the regulatory regions will vary, including regions from viral, plasmid or chromosomal genes, or the like. For expression in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or regulatable promoters may be employed. Expression in a microorganism can provide a ready source of the plant enzyme. Among transcriptional initiation regions which have been described are regions from bacterial and yeast hosts, such as *E. coli*, *B. subtilis*, *Sacchromyces cerevisiae*, including genes such as beta-galactosidase, T7 polymerase, tryptophan E and the like.

For the most part, the constructs will involve regulatory regions functional in plants. The open reading frame, coding for the plant vde or functional fragment thereof will be joined at its 5' end to a transcription initiation regulatory region such as the wild-type sequence naturally found 5' upstream to the vde structural gene. Numerous other transcription initiation regions are available which provide for a wide variety of constitutive or regulatable, e.g., inducible, transcription of the structural gene functions. Constitutive promoters such as the CaMV 35S promoter, double 35S promoter, 34S figwort promoter may be useful. Promoters which express in plastid containing cells will be of special interest. Some such promoters are preferentially expressed in plastid containing tissues, such as the ssu promoter. The transcription/translation initiation regions corresponding to such structural genes are found immediately 5' upstream to the respective start codons. In embodiments wherein the expression of the vde protein is desired in a plant host, the use of all or part of the complete plant vde gene is desired; namely all or part of the 5' upstream non-coding regions (promoter) together with the structural gene sequence and 3' downstream non-coding regions may be employed. If a different promoter is desired, such as a promoter native to the plant host of interest or a modified promoter, i.e., having

transcription initiation regions derived from one gene source and translation initiation regions derived from a different gene source, including the sequence encoding the plant vde of interest, or enhanced promoters, such as double 35S CaMV promoters, the sequences may be joined together using standard techniques.

Expression of the vde transcript was followed in market romaine lettuce leaves that were dark adapted for an undetermined period of time. The same level of transcript was detected in both yellow leaves and rapidly expanding green leaves. However, a greater transcript level was detected in mature green leaves. Two hybridizing transcripts were observed for each sample indicating the possibility that the upper larger transcript may be processed to the slightly smaller transcript (1.7 kb) having the greater level of hybridization. The increased level of transcript in mature green leaves of lettuce may be due to two possible reasons: higher expression occurs in tissues with a higher density of fully developed chloroplasts or expression may be regulated by light intensity since the mature green leaves receive a higher intensity of light than the immature leaves which are partially shielded in the center of the head of lettuce. Hence, use of the vde promoter may be particularly useful in the transcription of vde nucleic acid sequences or for the expression of other nucleic acid sequences of interest.

Regulatory transcript termination regions may be provided in DNA constructs of this invention as well. Transcript termination regions may be provided by the DNA sequence encoding the plant vde or a convenient transcription termination region derived from a different gene source, for example, the transcript termination region which is naturally associated with the transcript initiation region. Where the transcript termination region is from a different gene source, it will contain at least about 0.5 kb, preferably about 1-3 kb of sequence 3' to the structural gene from which the termination region is derived.

Plant expression or transcription constructs having a plant vde as the DNA sequence of interest for increased or decreased expression thereof may be employed with a wide variety of plant life, particularly, plant life where light regulation or zeaxanthin levels are important. Plants of interest include, but are not limited to ornamental plant varieties, field and forage crops, including alfalfa and trees. Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences may be required. Importantly, this invention is applicable to dicot and monocot species alike and will be readily applicable to new and/or improved transformation and regulation techniques.

The method of transformation in obtaining such transgenic plants is not critical to the instant invention, and various methods of plant transformation are currently available. Furthermore, as newer methods become available to transform crops, they may also be directly applied hereunder. For example, many plant species naturally susceptible to *Agrobacterium* infection may be successfully transformed via tripartite or binary vector methods of *Agrobacterium* mediated transformation. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders may find use with other modes of transformation. In addition, techniques of microinjection, DNA particle bombardment, and electroporation have been developed which allow for the transformation of various monocot and dicot plant species.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host, viral immunity or the like. Depending upon the number of

different host species the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be introduced into the *Agrobacterium* host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-plasmid containing the T-DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall formation), the latter being permissible, so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed plasmid can give a mixture of normal plant cells and gall.

In some instances where *Agrobacterium* is used as the vehicle for transforming host plant cells, the expression or transcription construct bordered by the T-DNA border region(s) will be inserted into a broad host range vector capable of replication in *E. coli* and *Agrobacterium*, there being broad host range vectors described in the literature. Commonly used is pRK2 or derivatives thereof. See, for example, Ditta, et al., (Proc. Nat. Acad. Sci., U.S.A. (1980) 77:7347-7351) and EPA 0 120 515, which are incorporated herein by reference. Alternatively, one may insert the sequences to be expressed in plant cells into a vector containing separate replication sequences, one of which stabilizes the vector in *E. coli*, and the other in *Agrobacterium*. See, for example, McBride and Summerfelt (Plant Mol. Biol. (1990) 14:269-276), wherein the pRiHRI (Jouanin, et al., Mol. Gen. Genet. (1985) 201:370-374) origin of replication is utilized and provides for added stability of the plant expression vectors in host *Agrobacterium* cells.

Included with the expression construct and the T-DNA will be one or more markers, which allow for selection of transformed *Agrobacterium* and transformed plant cells. A number of markers have been developed for use with plant cells, such as resistance to chloramphenicol, kanamycin, the

aminoglycoside G418, hygromycin, or the like. The particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

EXAMPLES

Example 1 - Lettuce vde cDNA

Vde was purified from romaine lettuce (*Lactuca sativa* L. cv Romaine) chloroplasts and peptides from a tryptic digest along with the N-terminus were sequenced (Rockholm, *Plant Physiol.* (1996) 110:697-703). Two peptides (N-terminus and tryptic fragment #15, shown in Fig.1) were used to develop the oligonucleotides

5'-GAYGCHYTBAAGACHTGYGC-3' (216-fold degeneracy) and

5'-TTGVARRTTDGGRATRAT-3' (144-fold degeneracy).

The partial cDNA for vde was amplified by 35 cycles of polymerase chain reaction (PCR) containing 25 pmol of each primer and lettuce cDNA using an annealing temperature of 50°C. The PCR product was subcloned into pGEM-7Zf (Promega) by blunt-end cloning and sequenced. A cDNA library was constructed from poly(A)+ RNA isolated from a pooled sample of

various age romaine lettuce leaves using the Timesaver cDNA Synthesis Kit (Pharmacia) and ligated into lambda-ZAPII (Stratagene). A total of 2.5×10^5 recombinant plaques were screened with the PCR product labeled by random priming and positive clones were plaque purified followed by *in vivo* excision of the plasmid. The cDNAs were subcloned into the *NotI* site of pGEM-5Zf and both strands of cDNA were sequenced completely using an Applied Biosystems Model 373A automated sequencer. The Genbank accession number is U31462.

The vde cDNA encompasses an open reading frame encoding a 473 amino acid protein with a calculated *Mr* of 54,447. The deduced protein contains an 125 amino acid putative transit peptide for transport into the chloroplast lumen where the enzyme is localized (Hager, *Planta* (1969)89:224-243). This was verified by *in vitro* transcription/translation of two vde (vde1:-234 to 1526 bp and vde2:-65 to 1578 bp of Fig. 1) cDNAs which produced a 55 Kd product on a sodium dodecyl sulfate (SDS)-polyacrylamide gel. The N-terminus of the mature vde protein (amino acid #126) was determined by N-terminal sequencing of purified vde from romaine lettuce. Therefore, mature vde consists of a 348 amino acid protein with a calculated *Mr* of 39,929 and a calculated *pI* of 4.57.

The primary structure of the deduced mature vde exhibits some characteristic features. The protein is hydrophilic overall with 57.2% of the total amino acid residues having polar side chains. Three interesting domains were identified in the deduced mature vde including a cysteine rich domain, a lipocalin signature and a highly charged domain. In the first domain 11 of the 13 total cysteines in the mature vde are present suggesting that this is most likely the site where dithiothreitol (DTT), a known inhibitor of vde, has its effect. The cysteines probably form more than one disulfide linkage since partial inhibition of vde activity with DTT results in an accumulation of antheraxanthin. The deduced mature vde also contains a lipocalin signature, a domain identified in a number of diverse proteins that bind small hydrophobic molecules. For example, crustacyanin, a protein

from lobster carapace which contains a lipocalin signature, binds the carotenoid astaxanthin. Similarly, this domain may play a role in binding the substrate violaxanthin. In the third domain approximately 47% of the residues have charged side chains. The most striking feature of this domain is the high concentration of glutamic acid residues; 27.6% of the residues in this domain (69.2% of the total in the mature vde) are glutamic acids whereas only 2% are aspartic acids

Figure 4 provides a detailed analysis of the deduced amino acid sequence of vde. The top portion provides a comparison of the deduced amino acid sequences of vde from three plant species. The transit peptides are located in the boxed region. Identical residues are indicated by hyphens (-). Gaps introduced to maximize sequence alignment are indicated by periods (.). Asterisks (*) identify the 13 cysteine residues that are conserved between the three sequences. ,

The bottom map of Figure 4 shows the three domains identified. The amino acid spanning regions for lettuce vde are indicated below the domains.

Figure 6 provides hydropathy profiles for the vdes from three species.

Example 2 - Expression of Lettuce vde cDNA in E.coli

Authenticity of the lettuce vde cDNA was confirmed by expression of the fragment vde2 in *E. coli*. Vde2 was subcloned in both sense and antisense orientations with respect to *lacZ* into the *NotI* site of pGEM-5Zf and transformed into *E. coli* DH5alpha. All cultures were incubated and induced with 10 mM IPTG (Bugos, *Plant Mol Biol.*(1991)17:1203-1215). Following the 2 hr induction, the cells were centrifuged at 4000xg for 10 min at 4°C. The cells were resuspended in 3 ml 50 mM Tris (pH 7.4), 1 mM EDTA and lysed using an ultrasonic cell disrupter equipped with a micro-probe for 10 cycles (30 sec on/30 sec off) while being cooled in an ice bath. The resulting extract was centrifuged at 1 0,000xg for 10 min at 4°C and the supernatant was collected for determining vde

activity using the *in vitro* assay and absorbance change at 502nm minus 540nm (Yamamoto, *Methods Enzymol.* (1985)110:303-312). The pellet was washed with 3 ml 50 mM Tris (pH 7.4), 1 mM EDTA and centrifuged. The pellet was resuspended in 3 ml buffer and assayed. All assays contained 100 μ l *E. coli* extract or pellet resuspension. For quantification of xanthophyll pigments, the reactions were stopped at various times with addition of solid Tris and the xanthophylls were extracted 3 times with diethyl ether. The ether was dried under a stream of N₂ and the xanthophylls were solubilized in 100 μ l 90% acetone followed by HPLC analysis (Gilmore, *J. Chromatogr.* (1991)543:137-145).

Extracts from *E. coli* expressing the fragment orientated with *lacZ* (sense) had strong vde activity whereas no detectable activity was observed from extracts of *E. coli* transformed with vde2 in antisense orientation or pGEM-5Zf alone. Furthermore, addition of DTT, a strong inhibitor of de-epoxidase activity, abolished all vde activity. DTT (3mM, final conc.) was added directly to the assay 50 seconds after ascorbate (30mM, final conc.) addition. Specific activity of the enzyme was 64.9 ± 5.4 nmols violaxanthin deepoxidized/min/mg protein. Trace activity was detected in the membrane fraction of vde2 sense suggesting that some of the enzyme was not washed away following lysis or that lysis was not complete. An attempt to express the vde1 fragment was unsuccessful. *E. coli* transformed with vde1 subcloned in pGEM-5Zf and orientated with *lacZ* did not grow.

To verify the products of de-epoxidation, the reaction with vde2 sense extract was stopped at various times and the xanthophylls were analyzed by HPLC. Antheraxanthin and zeaxanthin appeared consistent with sequential de-epoxidation and concomitant with the rapid decrease in violaxanthin, similar to observations reported over three decades earlier for de-epoxidation in lima bean (*Phaseolus leunatus*) leaves exposed to high light (Yamamoto, *Arch. Biochem. Biophys.* (1962)97:168-173). The specific activity of the enzyme was 19.4 ± 0.9 nmols violaxanthin de-epoxidized/min/mg protein. This

is the first unequivocal evidence that the same enzyme catalyzes the two-step mono de-epoxidation reaction.

Example 3 - vde in Other Plants

Western analysis of vde from chloroplasts of various C_3 plants and expressed vde in *E. coli* demonstrate that the N-terminus is conserved.

Intact chloroplasts were isolated (Neubauer, *Plant Physiol.* (1992)99:1354-1361) and lysed with five freeze/thaw cycles using liquid N_2 (Hager, *Planta* (1975)88:27-44). Expression of vde2 in *E. coli* DH5-alpha was as described in Example 2 and the cells were lysed using the freeze/thaw method. Proteins were resolved on a 12% SDS-polyacrylamide gel and electrophoretically transferred to PVDF. Color development was performed following incubation with alkaline phosphatase-conjugated secondary antibodies. Protein was estimated using a prepared reagent (Biorad) and bovine gamma globulin as the standard.

The blot was probed with a polyclonal antibody prepared against a synthetic peptide derived from the N-terminus of lettuce vde (VDALKTCACLLK). Vde migrates with an approximate size of 43 kD.

The mature vde from market romaine lettuce, tobacco (*Nicotiana tabacum* L. cv Xanthi) and market spinach (*Spinacia oleracea* L.) all migrate with approximately the same M_r of 43K. The antibody recognized vde in these three plant species demonstrating that the N-terminus is conserved. Expressed vde2 in *E. coli* migrated at the same M_r as the romaine lettuce vde whereas extracts from *E. coli* containing only pGEM-5Zf produce some minor cross-reacting proteins, none of which having a M_r of 43K. The M_r 's of the above vde proteins are in agreement with the calculated M_r of the deduced mature vde (39.9K). Two interesting observations are evident from vde expressed in *E. coli*. The first is that the *E. coli* expressed vde produced many immunoreactive bands of lower molecular weight. Reasons for this may be due to some processing occurring at the C-terminus of the protein by *E. coli* (since

the antibody recognizes the N-terminus) or due to translational pausing. The second is that the bacterial expressed vde protein migrates at the same molecular weight as mature vde from romaine lettuce and not as the expected size of the deduced vde preprotein (54.4K) with the transit peptide. This suggests that *E. coli* may recognize the chloroplast transit peptide and cleave it. The N-terminus of the bacterial expressed vde will need to be sequenced to determine the actual site where cleavage is occurring. A similar observation was also reported for the nuclear-encoded chloroplast enzyme acetolactate synthase from *Arabidopsis* when expressed in *E. coli*.

Figure 7 shows the kinetics of absorbance change demonstrating expression of active violaxanthin de-epoxidase in *E. coli* DH5 (top of Fig. 7). Expression was assayed from vde2 constructs in both sense and antisense orientations with respect to *lacZ* along with *E. coli* containing the vector only (pGEM-5Zf). DTT (3mM, final concentration) was added directly to the assay 70 seconds after ascorbate (30 mM, final concentration) addition. Specific activity of the enzyme was 64.9 ± 5.4 nmols violaxanthin de-epoxidized min⁻¹ mg. protein⁻¹.

The bottom of Figure 7 is a timecourse of xanthophyll conversions by expressed vde2 (sense construct) in *E. coli*. Specific activity of the enzyme was 19.4 ± 0.9 nmols violaxanthin de-epoxidized min⁻¹ protein⁻¹.

Example 4 - Effects of Expression of vde in Plants

In Figure 8, pigment analysis of leaves of 212 control tobacco plants (Ct-#) is provided, as well as the mean percentage of violaxanthin which is de-epoxidized. Also provided by Figure 8 is the pigment analysis of leaves of 18 vde-antisense tobacco plants (TAS-#).

Tobacco plants were transformed with an antisense construct of the tobacco vde cDNA under control of the CaMV 35S promoter (pB1121) using *Agrobacterium tumefaciens* LBA4404.

A total of 40 antisense plants were analyzed with 18 showing various levels of inhibition of de-epoxidation.

Relative pigment concentration for tobacco (*Nicotiana tabacum* L. cv. Xanthi) leaves was measured by leaf disks punched from tobacco leaves that were dark adapted for a few hours. One leaf disk (dark adapted) was extracted with acetone and analyzed by HPLC while another was light induced by exposing the disk to $1800 \text{ } \mu\text{mol m}^{-2} \text{ s}^{-1}$ white light for 20 min while the leaf disk was floating on water in a water-jacketed beaker cooled at 20°C. Following the light treatment, the leaf disk was extracted and analyzed by HPLC.

Two vde-antisense tobacco plants (TAS-32 and TAS-39) were recovered that had undetectable levels of zeaxanthin following illumination with bright white light. Low levels of antheraxanthin (~2-3%) were present in some dark-adapted leaves and are assumed to represent incomplete epoxidase activity.

In Figures 9 and 10, results are provided from a comparison of measurements on a tobacco leaf from a control plant (Ct-30) and a vde-antisense plant (TAS-5), both of which were dark adapted for 24 hours. Under low light conditions, three leaf disks were punched from each leaf. One leaf disk (dark adapted) was extracted and analyzed by HPLC.

The remaining two leaf disks were pre-illuminated with $500 \text{ } \mu\text{mol m}^{-2} \text{ s}^{-1}$ red light for 15 minutes. One of these disks was then extracted and analyzed by HPLC while the other was placed in the dark for 10 minutes prior to fluorometry and HPLC analysis.

It has also been observed that in tobacco plants where lettuce vde has been overexpressed from a 35S construct, flowering is delayed, and flowers are slightly larger.